

CONTENTS

0 Random Variables and Their Distributions

0.1 Notation and Terminology	4
0.2 Discrete Random Variables	7
0.3 Continuous Random Variables	14
0.4 Properties of Random Variables	16
0.5 Some Properties of Expected Values	21
0.6 Moment Generating Functions	26
0.7 Probability Generating Function	30
0.8 Cumulant Generating Function	31
0.9 Special Discrete Distributions	32
0.9.1 Bernoulli Distribution	32
0.9.2 Binomial Distribution	33
0.9.3 Hypergeometric Distribution	35
0.9.4 Geometric Distributions	37
0.9.5 Negative Binomial	40
0.9.6 Poisson Distribution	42

0.9.7 Discrete Uniform Distribution	44
0.10 Special Continuous Distributions	45
0.10.1 Uniform Distribution	45
0.10.2 Gamma Distribution	46
0.10.3 Exponential Distribution	51
0.10.4 Weibull Distribution	53
0.10.5 Pareto Distribution	55
0.10.6 Normal Distribution	56
0.10.7 Log Normal Distribution	58
0.10.8 Beta Distribution	59
0.11 Location and Scale Parameters	61
0.11.1 Cauchy Distribution	64
0.11.2 Two-parameter Exponential Distribution	65
0.11.3 Double-Exponential Distribution	66
0.12 Joint Discrete Distributions	66
0.13 Joint Continuous Distributions	75
0.14 Conditional Distributions	82
0.15 Independent Random Variables	88

Chapter 0 Random Variables and Their Distributions-Review 1	3
0.16 The Expected Value of a Function of Random Variables	94

0 Random Variables and Their Distributions	94
--	----

0.1 Notation and Terminology

- **Experiment** refers to the process of obtaining an observed result of some phenomenon. It could pertain to activities as scientific experiments or games of chance.
- **Trial** of the experiment is a performance of an experiment.
- The set of all possible outcomes of an experiment is called the **sample space**, denoted by S .
- If a sample space S is either finite or countably infinite then it is called a **discrete sample space**.
- An **event** is a subset of the sample space S . If A is an event, then A has occurred if it contains the outcome that occurred.
- **Random variable**, say X , is a function de-

fined over a sample space, S , that associates a real number, $X(e) = x$, with each possible outcome e in S .

Example 1. An experiment consists of tossing two coins, and the observed face of each coin is of interest. The sample space is

Sol:

$$S = \{HH, HT, TH, TT\}$$

Example 2. Suppose that in Example 23 we were not interested in the individual outcomes of the coins, but only in the total number of heads obtained from the two coins. An appropriate sample space is

Sol:

$$S = \{0, 1, 2\}$$

Thus, different sample spaces may be appropriate for the same experiment, depending on the characteristic of interest.

0.2 Discrete Random Variables

Definition 1. If the set of all possible values of a random variable, X , is a countable set, x_1, x_2, \dots, x_n , or x_1, x_2, \dots , then X is called a discrete random variable. The function

$$f(x) = P[X = x] \text{ for } x = x_1, x_2, \dots$$

that assigns the probability to each possible value x will be called the **discrete probability density function** (discrete pdf).

Theorem 1. A function $f(x)$ is a discrete pdf if and only if it satisfies both of the following properties for at most a countably infinite set of reals x_1, x_2, \dots .

$$f(x_i) \geq 0 \text{ for all } x_i,$$

and

$$\sum_{\text{all } x_i} x_i = 1$$

In some problems, it is possible to express the pdf by means of an equation. However, it is sometimes more convenient to express it in tabular form.

Note: If it is clear from the context that X is discrete, then we simply will say pdf. Another common terminology is probability mass function (pmf), and the possible values, x , are called mass points of X . Sometimes a subscripted notation, $f_X(x)$, is used.

Example 5. We roll a red die and a green die.

Both dice are fair. Suppose X is the total score from the red and green dice.

(a) What are the possible values of X ?

Sol:

$$x = 2, 3, \dots, 12$$

Here, the set of the possible values of X is finite

(b) Display the distribution of X in a table.

Sol:

x	2	3	4	5	6	7	8	9	10	11	12
$P(X = x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

$$\begin{aligned} P(X = 1) &= p \\ P(X = 2) &= qp \\ P(X = 3) &= q^2 p \\ &\vdots \\ P(X = x) &= q^{x-1} p \end{aligned}$$

Example 7.

If we roll a 12-sided die twice. If each face is marked with an integer, 1 through 12, then each value is equally likely to occur on a single roll of the die. Let X be the maximum obtained on the two rolls. Find the pdf of X .

Sol:

It is not hard to see that for each value x there are an odd number, $2x - 1$, of ways for that value to occur. Thus, the pdf of X must have the form $f(x) = c(2x - 1)$ for $x = 1, 2, \dots, 12$

$$\sum_{j=1}^{12} f(x_j) = 1$$

$$c \sum_{j=1}^{12} (2x_j - 1) = 1$$

$$c \left[\frac{2(12)(1+12)}{12} - 12 \right] = 1$$

$$c = \frac{1}{144}$$

Definition 2. The cumulative distribution function (CDF) of a random variable X is defined for any real x by

$$F(x) = P[X \leq x]$$

$$f(x_i) = F(x_i) - F(x_{i-1})$$

Furthermore, if $x < x_1$ then $F(x) = 0$, and for any other real x

$$F(x) = \sum_{x_i \leq x} f(x_i)$$

where the summation is taken over all indices i such that $x_i \leq x$.

The CDF of any random variable must satisfy the properties of the following theorem.

Theorem 3.

A function $F(x)$ is a CDF for some random variable X if and only if it satisfies the following properties:

- $\lim_{x \rightarrow -\infty} F(x) = 0$
- $\lim_{x \rightarrow \infty} F(x) = 1$
- $\lim_{h \rightarrow 0^+} F(x + h) = F(x)$, i.e. $F(x)$ is **continuous from the right**
- $a < b$ implies $F(a) \leq F(b)$, i.e. $F(x)$ is **non-decreasing**

The first two properties say that $F(x)$ can be made **arbitrarily** close to 0 or 1 by taking x arbitrarily large, and negative and positive, respectively.

0.3 Continuous Random Variables

Definition 3. A random variable X is called a continuous random variable if there is a function $f(x)$, called the probability density function (pdf) of X , such that the CDF can be represented as

$$F(x) = \int_{-\infty}^x f(t)dt$$

Theorem 4. A function $f(x)$ is a pdf for some continuous random variable X if and only if it satisfies the properties

$$f(x) \geq 0 \text{ for all real } x,$$

and

$$\int_{-\infty}^{\infty} f(x)dx = 1$$

Example 8.

A machine produces copper wire, and occasionally there is a flaw at some point along the wire. The length of wire (in meters) produced between successive flaws is a continuous random variable X with pdf of the form

$$f(x) = \begin{cases} c(1+x)^{-3} & x > 0 \\ 0 & \text{otherwise} \end{cases}$$

where c is a constant. Find c and hence $F(x)$.

Sol:

$$\int_0^\infty c(1+x)^{-2} dx = 1$$

$$\frac{1}{2}c = 1$$

$$c = 2$$

$$\begin{aligned} F(x) &= \int_0^x \frac{1}{2}(1+t)^{-3} dt \\ &= 2 \frac{(1+t)^{-2}}{-2} \Big|_0^x \\ &= 1 - (1+x)^{-2} \end{aligned}$$

0.4 Properties of Random Variables

Definition 4. If X is a discrete random variable with pdf $f(x)$, then the expected value of X is defined by

$$E(X) = \sum_x x f(x)$$

Definition 5. If X is a continuous random variable with pdf $f(x)$, then the expected value of X is defined by

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx$$

if the integral is absolutely convergent. Otherwise we say that $E(X)$ does not exist.

Other notations for $E(X)$ are μ or μ_X , and the terms mean or expectation of X also are commonly used.

Definition 6. If $0 < p < 1$, then a $100 \times p^{th}$ percentile of the distribution of a continuous random variable X is a solution x to the equation

$$F(x) = p$$

In general, a distribution may not be continuous, and if it has a discontinuity, then there will be some values of p for which equation $F(x) = p$ has no solution. It is possible to state a general definition of percentile by defining a p^{th} percentile of the distribution of X to be a value x_p , such that $P[X \leq x_p] \geq p$ and $P[X \geq x_p] \leq 1 - p$. In essence, x_p is a value such that $100 \times p$ percent of the population values are at most x_p , and $100 \times (1 - p)$ percent of the population values are at least x_p . A median of the distribution of X is a 50-th percentile, denoted by $x_{0.5}$ or m .

Example 9. A discrete random variable X has a pdf of the form $f(x) = c(8 - x)$ for $x = 0, 1, 2, 3, 4, 5$, and zero otherwise. Find $E(X)$.

Sol:

$$c(8 + 7 + 6 + 5 + 4 + 3) = 1$$

$$c = \frac{1}{33}$$

$$E(X) = \frac{1}{33}(1 \times 7 + 2 \times 6 + 3 \times 5 + 4 \times 4 + 5 \times 3) = \frac{65}{33}$$

Example 10.

A continuous random variable X has a pdf of the form $f(x) = \frac{2x}{9}$ for $0 < x < 3$, and zero otherwise.

- Find a number m such that $P[X \leq m] = P[X \geq m]$.

Sol:

$$\begin{aligned} P[X \leq m] &= P[X \geq m] \\ \int_0^m \frac{2x}{9} dx &= \int_m^3 \frac{2x}{9} dx \\ \frac{x^2}{9} \Big|_0^m &= \frac{x^2}{9} \Big|_m^3 \\ \frac{m^2}{9} &= 1 - \frac{m^2}{9} \\ \frac{2m^2}{9} &= 1 \end{aligned}$$

$$m = \frac{3}{\sqrt{2}} = \frac{3\sqrt{2}}{2}$$

- Find $E(X)$.

Sol:

$$E(X) = \int_0^3 \frac{2x^2}{9} dx = \frac{2x^3}{27} \Big|_0^3 = 2$$

Example 11. Consider the distribution of lifetimes, X (in months), of a particular type of component. We will assume that the CDF has the form

$$F(x) = 1 - e^{-(x/3)^2}, X > 0$$

and zero otherwise.

- Find the median lifetime.

Sol:

$$\begin{aligned} F(m) &= \frac{1}{2} \\ 1 - e^{-(m/3)^2} &= \frac{1}{2} \\ -\frac{m^2}{9} &= \ln \frac{1}{2} \\ m^2 &= 9 \ln 2 \\ m &= 3\sqrt{\ln 2} = 2.498 \text{ months} \end{aligned}$$

- Find the time t such that 10% of the components fail before t .

Sol:

$$\begin{aligned} F(t) &= 0.1 \\ 1 - e^{-(t/3)^2} &= 0.1 \\ -\frac{t^2}{9} &= \ln 0.9 \\ t &= 3\sqrt{-\ln 0.9} = 0.974 \text{ months} \end{aligned}$$

0.5 Some Properties of Expected Values

Theorem 5. If X is a random variable with pdf $f(x)$ and $u(x)$ is a real valued function whose domain includes the possible values of X , then

$$E[u(X)] = \sum u(x) f(x) \text{ if } X \text{ is discrete}$$

$$E[u(X)] = \int_{-\infty}^{\infty} u(x) f(x) dx \text{ if } X \text{ is continuous}$$

Theorem 6. If X is a random variable with pdf $f(x)$, a and b are constants, and $g(x)$ and $h(x)$ are real valued functions whose domains include the possible values of X , then

$$E[ag(X) + bh(X)] = aE[g(X)] + bE[h(X)]$$

and the k^{th} moment about the mean is

$$\mu'_k = E(X^k)$$

Other common notations for the variance are σ^2 , σ_X^2 , or $V(X)$, and a related quantity, called the standard deviation of X , is the positive square root of the variance, $\sigma = \sqrt{V(X)}$. The variance provides a measure of the variability or amount of “spread” in the distribution of a random variable.

Definition 8. The k^{th} moment about the origin of a random variable X is

$$\mu_k' = E(X^k)$$

Definition 7. The variance of a random variable X is given by

$$V(X) = E[(X - \mu)^2]$$

Theorem 7. If X is a random variable, then

$$V(X) = E(X^2) - \mu^2$$

Sol:

$$\begin{aligned} V(X) &= E(X - \mu)^2 \\ &= E(X^2 - 2\mu X + \mu^2) \\ &= E(X^2) - 2\mu E(X) + \mu^2 \\ &= E(X^2) - 2\mu^2 + \mu^2 \\ &= E(X^2) - \mu^2 \end{aligned}$$

(a) Find $E(X)$.

Theorem 8. If X is a random variable and a and b are constants, then

$$V(aX + b) = a^2 V(X)$$

Sol:

$$\begin{aligned} V(aX + b) &= E[aX + b - (a\mu + b)]^2 \\ &= E[aX - a\mu]^2 \\ &= a^2 E(X - \mu)^2 \\ &= a^2 V(X) \end{aligned}$$

This means that the variance is affected by a change of scale, but not by a translation.

Example 13. Let X be continuous with pdf
 $f(x) = 3x^2$ if $0 < x < 1$ and zero otherwise.
 Find

(a) $E(X)$.

Sol:

$$E(X) = \int_0^1 3x^3 dx = \frac{3x^4}{4} \Big|_0^1 = \frac{3}{4}$$

(b) $V(X)$

Sol:

$$\begin{aligned} E(X^2) &= \int_0^1 3x^4 dx = \frac{3x^5}{5} \Big|_0^1 = \frac{3}{5} \\ V(X) &= \frac{3}{5} - \frac{9}{16} = 0.1875 \end{aligned}$$

(c) $E(X^r)$

Sol:

$$E(X^r) = \int_0^1 3x^{2+r} dx = \frac{3x^{3+r}}{3+r} \Big|_0^1 = \frac{3}{3+r}$$

(d) $E(3X - 5X^2 + 1)$

Sol:

$$\begin{aligned} E(3X - 5X^2 + 1) &= 3E(X) - 5E(X^2) + 1 = \\ 3\left(\frac{3}{4}\right) - 5\left(\frac{3}{5}\right) + 1 &= 0.25 \end{aligned}$$

0.6 Moment Generating Functions

Definition 9. If X is a random variable, then
 the expected value

$$(a) M_X(t) = E(e^{tX})$$

Sol:
 is called the moment generating function (MGF)
 of X if this expected value exists for all values
 of t in some interval of the form $-h < t < h$ for
 some $h > 0$.

Theorem 9. If the MGF of X exists, then

$$(b) E(X^r) = M^r(0) \text{ for all } r = 1, 2, \dots$$

Example 14. Consider a continuous random variable X with pdf $f(x) = e^{-x}$ if $x > 0$ and zero otherwise.

(a) Find the MGF of X .

(b) Find $E(X^r)$.

(c) Find the mean and variance of X

Sol:

$$\begin{aligned} \text{(a)} \quad M_X(t) &= E[e^{tX}] = \int_0^\infty e^{tx} e^{-x} dx \\ &= \int_0^\infty e^{-(1-t)x} dx = \frac{-e^{-(1-t)}}{1-t} \Big|_0^\infty \\ &= \frac{1}{1-t}, t < 1 \end{aligned}$$

$$\begin{aligned} \text{(b)} \quad M'_X(t) &= \frac{1}{(1-t)^2} \\ M''_x(t) &= \frac{2}{(1-t)^3} \\ \vdots & \end{aligned}$$

$$\begin{aligned} M_X^r(t) &= \frac{r!}{(1-t)^{r+1}} \\ E(X^r) &= M_X^r(0) = r! \end{aligned}$$

$$\begin{aligned} \text{(c)} \quad E(X) &= 1 \\ E(X^2) &= 2! = 2 \\ V(X) &= 2 - 1 = 1 \end{aligned}$$

Properties of Moment Generating Functions

Theorem 10. If $Y = aX + b$, then $M_Y(t) = e^{bt}M_X(at)$.

Sol:

$$\begin{aligned} M_Y(t) &= E[e^{(aX+b)t}] \\ &= e^{bt}E[e^{(at)X}] \\ &= e^{bt}M_X(at) \end{aligned}$$

Theorem 11. Uniqueness If X_1 and X_2 have respective CDFs $F_1(x)$ and $F_2(x)$, and MGFs $M_1(t)$ and $M_2(t)$, then $F_1(x) = F_2(x)$ for all real x if and only if $M_1(t) = M_2(t)$ for all t in some interval $-h < t < h$ for some $h > 0$

In other words, X_1 and X_2 cannot have the same MGF but different pdf's. Thus, the form of the MGF determines the form of the pdf.

0.8 Cumulant Generating Function

The cumulant-generating function $K(t)$, is the natural logarithm of the moment-generating function:

$$K(t) = \ln E(e^{tX}) = \ln M_X(t)$$

The cumulants κ_n are obtained from a power series expansion of the cumulant generating function:

$$K(t) = \sum_{n=1}^{\infty} \kappa_n \frac{t^n}{n!}$$

This expansion is a Maclaurin series, so that n^{th} cumulant can be obtained by differentiating the above expansion n times and evaluating the result at zero:

$$\kappa_n = K^{(n)}(0).$$

The first cumulant is the expected value; the second and third cumulants are respectively the second and third central moments; but the higher cumulants are neither moments nor central moments.

0.9.2 Binomial Distribution

A random variable X has Binomial (n, p) distribution if its pdf is

$$f(x) = P(X = x) = \binom{n}{x} p^x q^{n-x} \text{ for } x = 0, 1, \dots, n$$

The Binomial (n, p) distribution arises as follows. Repeat a Bernoulli experiment independently n times and each time one observes the outcome 0 or 1 and $p = P(X = 1)$.

A short notation to designate that X has the binomial distribution with parameters n and p is $X \sim BIN(n, p)$

The mean, variance and MGF of a Binomial distribution are:

$$E(X) = np, \sigma^2 = np(1-p), M_X(t) = [pe^t + q]^n$$

Notes:

- $x(n) = n(n-1), (n) = \frac{N}{n}(N-1)$
- Binomial Theorem: $(x+y)^n = \sum_{i=1}^n \binom{n}{i} x^i y^{n-i}$

Example 16.

In a 10-question true/false test:

- (a) What is the probability of getting all answers correct by guessing?
- (b) What is the probability of getting eight correct by guessing?

Sol:

Let X be the number of questions answer correctly. Then $X \sim Bin(n = 10, p = 0.5)$

$$\begin{aligned} P(X = 10) &= 0.5^{10} \\ P(X = 8) &= \binom{10}{8} (0.5^8)(0.5^2) \end{aligned}$$

0.9.3 Hypergeometric Distribution

Suppose a population or collection consists of a finite number of items, say N , and there are M items of type 1 and the remaining $N - M$ items are of type 2. Suppose n items are drawn at random without replacement, and denote by X the number of items of type 1 that are drawn. The random variable X is said to have the hypergeometric distribution with parameters N , n and M . Its pdf is

$$f(x) = \frac{\binom{M}{x} \binom{N-M}{n-x}}{\binom{N}{n}},$$

$$x = 0, 1, \dots, \min(n, M), n - x \leq N - M.$$

We write $X \sim Hyp(n, M, N)$.

The mean and variance of a Hypergeometric distribution are:

$$E(X) = \frac{nM}{N}, \sigma^2 = n \frac{M}{N} \left(1 - \frac{M}{N}\right) \frac{N-n}{N-1}$$

The MGF of Hypergeometric distribution does not exists.

0.9.4 Geometric Distributions

If we denote the number of trials required to obtain the first success by X , then X is said to have **Geometric distributions**, the discrete pdf of X is given by

$$f(x) = pq^{x-1} \quad x = 1, 2, 3, \dots$$

We denote $X \sim Geo(p)$
The CDF of X is

$$F(x) = 1 - q^x \quad x = 1, 2, 3, \dots$$

Example 18. A geological exploration may indicate that a well drilled for oil in a region in Texas would strike oil with probability 0.3. Assuming that oil strikes are independent from one drill to another. What is the probability that the first oil strike will occur on the 6th drill?

Sol:

Let X be the number of well drill require to obtain the first oil strike.

$$\begin{aligned} X &\sim Geo(0.3) \\ P(X = 6) &= 0.3(0.7^6) = \boxed{0.050421} \end{aligned}$$

Theorem 12. No-Memory Property If

$$X \sim GEO(p),$$

then

$$P[X > j + k | X > j] = P[X > k]$$

Thus, knowing that j trials have passed without a success does not affect the probability of k more trials being required to obtain a success. That is, having several failures in a row does not mean that you are more “due” for a success.

0.9.5 Negative Binomial

In repeated independent Bernoulli trials, let X denote the number of trials required to obtain r successes. Then the probability distribution of X is the negative binomial distribution with discrete pdf given by

$$f(x) = \binom{x-1}{r-1} p^r q^{x-r}, x = r, r+1, \dots$$

A special notation, which designates that X has the negative binomial distribution

$$X \sim NB(r, p)$$

The mean, variance and MGF of a Negative Binomial distribution are:

$$E(X) = \frac{1}{p}, \sigma^2 = \frac{q}{p^2}, M_X(t) = \frac{pe^t}{1 - qe^t}$$

$$E(X) = \frac{r}{p}, \sigma^2 = \frac{rq}{p^2}, M_X(t) = \left(\frac{pe^t}{1 - qe^t} \right)^r$$

The mean, variance and MGF of a Negative Binomial distribution are:

Example 19. Team A plays team B in a seven-game world series. That is, the series is over when either team wins four games. For each game, $P(A \text{ wins}) = 0.6$, and the games are assumed independent. What is the probability that the series will end in exactly six games?

Sol: Let X and Y be the number of games play until team A and team B wins 4 games respectively.

$$X \sim NB(r = 4, p = 0.6), Y \sim NB(r = 4, p = 0.4)$$

$$\begin{aligned} P[\text{Team A wins 4 games}] \\ = P[X = 6] \\ = \binom{6-1}{4-1} 0.6^4 (0.4^{6-4}) \\ = 0.20763 \end{aligned}$$

$$\begin{aligned} P[\text{Team B wins 4 games}] \\ = P[Y = 6] \\ = \binom{6-1}{4-1} 0.4^4 (0.4^{6-4}) \\ = 0.09216 \end{aligned}$$

$$P[\text{series end in six games}] = 0.20763 + 0.09216 = \boxed{0.20736}$$

Example 20.

We are inspecting a particular brand of concrete slab specimens for visible cracks. Suppose that the number (X) of cracks per concrete slab has a Poisson distribution with $\mu = 2.5$. What is the probability that a randomly selected slab will have at least 2 cracks?

Sol:

$$\begin{aligned} X &\sim POI(2.5) \\ P[X \geq 2] &= 1 - P(X < 2) \\ &= 1 - P(X = 0) - P(X = 1) \\ &= 1 - e^{-2.5} - 2.5e^{-2.5} \\ &= \boxed{0.7127} \end{aligned}$$

0.9.7 Discrete Uniform Distribution

A discrete random variable X has the discrete uniform distribution on the integers $1, 2, \dots, N$ if it has a pdf of the form

$$f(x) = \frac{1}{N}, X = 1, 2, \dots, N$$

A special notation for this situation is

$$X \sim DU(N)$$

The mean, variance and MGF of a Discrete Uniform distribution are:

$$\begin{aligned} E(X) &= \frac{N+1}{2}, \sigma^2 = \frac{N^2 - 1}{12}, \\ M_X(t) &= \frac{1}{N} \frac{e^t - e^{(N+1)t}}{1 - e^t} \end{aligned}$$

Example:

- The number obtained by rolling an ordinary six-sided die correspond to $DU(6)$.
- The multiple-choice test on any question, which associate the four choices with the integers 1, 2, 3, and 4, then the response, X , on any given question that is answered at random is $DU(4)$.

0.10 Special Continuous Distributions

0.10.1 Uniform Distribution

A continuous random variable X that assume values only in a bounded interval (a, b) , with constant pdf over the interval is known as the **uniform distribution**.

$$f(x) = \frac{1}{b-a}, a < x < b$$

and zero otherwise. A notation that designates that X has pdf of the form above is

$$X \sim U(a, b)$$

The CDF of $X \sim U(a, b)$ has the form

$$F(x) = \begin{cases} 0, & x \leq a \\ \frac{x-a}{b-a}, & a < x < b \\ 1, & b \leq x \end{cases}$$

The mean, variance and MGF of a Discrete Uniform distribution are:

$$E(X) = \frac{a+b}{2}, \sigma^2 = \frac{(b-a)^2}{12},$$

$$M_X(t) = \frac{e^{bt} - e^{at}}{(b-a)t}$$

0.10.2 Gamma Distribution

Definition 10. The gamma function, denoted by $\Gamma(\alpha)$ for all $\alpha > 0$, is given by

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt$$

Theorem 13. The gamma function satisfies the following properties:

- $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$
- $\Gamma(n) = (n - 1)!$
- $\Gamma(\frac{1}{2}) = \sqrt{\pi}$

Sol:

- Use $\int u dv = uv - \int v du$
- Let $u = t^{\alpha-1}$, $du = (\alpha - 1)t^{\alpha-2}dt$, $dv = e^{-t}dt$, $v = \int e^{-t}dt = -e^{-t}$

$$\begin{aligned} \Gamma(\alpha) &= \int_0^\infty t^{\alpha-1} e^{-t} dt \\ &= t^{\alpha-1} e^{-t} \Big|_0^\infty + (\alpha - 1) \int_0^\infty t^{\alpha-2} e^{-t} dt \\ &= (\alpha - 1)\Gamma(\alpha - 1) \end{aligned}$$

- $\Gamma(n) = (n - 1)\Gamma(n - 1) = (n - 1)(n - 2)\Gamma(n - 2) \dots (n - 1)(n - 2) \dots 1\Gamma(1) = (n - 1)!$

$$\begin{aligned}\bullet \Gamma\left(\frac{1}{2}\right) &= \int_0^{\infty} t^{-\frac{1}{2}} e^{-t} st \\ &= \text{Let } u = t^{\frac{1}{2}}, du = \frac{1}{2}u^{-\frac{1}{2}}dt \\ &= \int_0^{\infty} u^{-1} e^{-u^2} 2udu \\ &= 2 \int_0^{\infty} e^{-u^2} du\end{aligned}$$

$$\begin{aligned}[\Gamma\left(\frac{1}{2}\right)]^2 &= \left[\Gamma\left(\frac{1}{2}\right)\right] \left[\Gamma\left(\frac{1}{2}\right)\right] \\ &= \left[2 \int_0^{\infty} e^{-u^2} du\right] \left[2 \int_0^{\infty} e^{-v^2} dv\right]\end{aligned}$$

Let $u = r\cos\theta$, $v = r\sin\theta$, then $0 \leq r \leq \infty$ and $0 \leq \theta \leq \frac{\pi}{2}$.

$$\begin{aligned}J &= \left| \frac{\partial u}{\partial r} \frac{\partial v}{\partial \theta} - \frac{\partial u}{\partial \theta} \frac{\partial v}{\partial r} \right| \\ &= \left| \begin{matrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\sin\theta \end{matrix} \right| \\ &= r\cos^2\theta + r\sin^2\theta \\ &= r\end{aligned}$$

- A continuous random variable X is said to have the **gamma distribution** with parameters $\theta > 0$ and α if it has pdf of the form

$$f(x) = \frac{1}{\theta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\theta}, x > 0$$

and zero otherwise.

A special notation, which designates that X has pdf given by equation above, is

$$X \sim GAM(\alpha, \theta)$$

The parameter α is called a shape parameter because it determines the basic shape of the graph of the pdf. θ is called the scaled parameter.

The mean, variance and MGF of a Gamma distribution are:

$$E(X) = \alpha\theta, \sigma^2 = \alpha\theta^2, M_X(t) = \left(\frac{1}{1-\theta t}\right)^\alpha$$

$$\begin{aligned}[\Gamma\left(\frac{1}{2}\right)]^2 &= 4 \int_0^{\pi/2} \int_0^{\infty} e^{-r^2} r dr d\theta \\ &= 4 \int_0^{\pi/2} d\theta \int_0^{\infty} r e^{-r^2} dr \\ &= 4(\pi/2) \left[-\frac{1}{2}e^{-r^2}\right]_0^{\infty} \\ &= \pi\end{aligned}$$

Finally, since $e^{-u^2} > 0$ for all $u > 0$, then

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$$

Theorem 14. If $X \sim GAM(n, \theta)$, where n is a positive integer, then the CDF can be written

$$F(x) = 1 - \sum_{i=0}^{n-1} \frac{(x/\theta)^i}{i!} e^{-x/\theta}$$

Example 21. The daily amount (in inches) of measurable precipitation in a river valley is a random variable $X \sim GAM(\alpha = 6, \theta = 0.2)$. Find the probability that the amount of precipitation will exceed 2 inches.

Sol:

$$\begin{aligned} P(X > 2) \\ = \sum_{i=0}^5 \frac{10^i e^{-10}}{i!} \\ = \boxed{0.067} \end{aligned}$$

0.10.3 Exponential Distribution

A continuous random variable X has the exponential distribution with parameter $\theta > 0$ if it has a pdf of the form

$$f(x) = \frac{1}{\theta} e^{-x/\theta}, x > 0$$

and zero otherwise.

The CDF of X is

$$F(x) = 1 - e^{-x/\theta}, x > 0$$

The notation $X \sim GAM(1, \theta)$ could be used to designate that X , but a more common notation is

$$X \sim EXP(\theta)$$

The mean, variance and MGF of an Exponential distribution are:

$$E(X) = \theta, \sigma^2 = \theta^2, M_X(t) = \left(\frac{1}{1 - \theta t} \right)$$

0.10.4 Weibull Distribution

A widely used continuous distribution is named after the physicist W. Weibull, who suggested its use for numerous applications, including fatigue and breaking strength of materials. It is also a very popular choice as a failure-time distribution. A continuous random variable X is said to have the **Weibull distribution** with parameters $\tau > 0$ and $\theta > 0$ if it has a pdf of the form

$$f(x) = \frac{\tau}{\theta^\tau} x^{\tau-1} e^{-(x/\theta)^\tau}, x > 0$$

and zero otherwise. A notation that designates that X is

$$X \sim WEI(\tau, \theta)$$

The CDF of X is

$$F(x) = 1 - e^{-(x/\theta)^\tau}$$

The mean and variance of a Weibull distribution are:

$$E(X) = \theta \Gamma\left(1 + \frac{1}{\tau}\right), \sigma^2 = \theta^2 \left[\Gamma\left(1 + \frac{2}{\tau}\right) - \Gamma^2\left(1 + \frac{1}{\tau}\right) \right]$$

The MGF does not exist.

0.10.5 Pareto Distribution

A continuous random variable X is said to have the Pareto distribution with parameters $\alpha > 0$ and $\theta > 0$ if it has a pdf of the form

$$f(x) = \frac{\alpha\theta^\alpha}{(x+\theta)^{\alpha+1}}, x > 0$$

and zero otherwise. A notation to designate that X is

$$X \sim PAR(\alpha, \theta)$$

The CDF is given by

$$F(x) = 1 - \left(\frac{\theta}{x+\theta}\right)^\alpha$$

The mean and variance of a Pareto distribution are:

$$E(X) = \frac{\theta}{\alpha-1}, \sigma^2 = \frac{\theta^2}{(\alpha-1)^2(\alpha-2)}$$

The MGF does not exist.

0.10.6 Normal Distribution

A random variable X follows the normal distribution with mean μ and variance σ^2 if it has the pdf

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2},$$

for $x \in \mathbb{R}$, $\mu \in \mathbb{R}$ and $\sigma > 0$. This denoted by

$$X \sim N(\mu, \sigma^2)$$

Let $Z = \frac{X-\mu}{\sigma}$, then $Z \sim N(0, 1)$. Z is called Standard Normal distribution. Its pdf is

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, -\infty < z < \infty$$

The standard normal CDF is given by

$$\Phi(x) = \int_{-\infty}^x \phi(t) dt$$

The CDF of X is

$$F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

The MGF of Normal distribution is

$$M_X(t) = e^{\mu t + \sigma^2/2}$$

Theorem 16. If $X \sim N(\mu, \sigma^2)$ then

1. $E(X - \mu)^{2r} = \frac{(2r)!\sigma^{2r}}{r!2^r}$
2. $E(X - \mu)^{2r-1} = 0$

0.10.7 Log Normal Distribution

A random variable X follows the lognormal distribution with parameters μ and σ if it has the pdf

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}}e^{(\ln x - \mu)^2/2\sigma^2},$$

for $x > 0$, $\mu > 0$ and $\sigma > 0$. This denoted by

$$X \sim LN(\mu, \sigma)$$

The CDF of X is

$$F(x) = \Phi\left(\frac{\ln x - \mu}{\sigma}\right)$$

The k^{th} raw moment of lognormal distribution is

$$E(X^k) = e^{k\mu + \frac{1}{2}k^2\sigma^2}$$

Note: If $X \sim N(\mu, \sigma^2)$, then $u = e^X \sim LN(\mu, \sigma)$

0.10.8 Beta Distribution

The beta family of distributions is a continuous family on $(0, 1)$ indexed by two parameters. The $\text{beta}(a, b)$ pdf is

$$f(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1} (1-x)^{b-1},$$

$$0 < x < 1, a > 0, b > 0.$$

The mean and variance of X are

$$E(X) = \frac{a}{a+b}$$

and

$$V(Y) = \frac{ab}{a+b+1}(a+b)^2$$

In order to show that the pdf of beta distributions sum to one, we need to find

$$B(a, b) = \int_0^1 x^{a-1} (1-x)^{b-1} dx,$$

where $B(a, b)$ is called the beta function. The beta function is related to the gamma function

through the following identity:

$$B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}.$$

Sol.

The Relationship between gamma function and beta function

$$\begin{aligned} & \Gamma(a)\Gamma(b) \\ &= \int_{u=0}^{\infty} u^{a-1} e^{-u} du \int_{v=0}^{\infty} v^{b-1} e^{-v} dv \\ &= \int_{v=0}^{\infty} \int_{u=0}^{\infty} u^{a-1} v^{b-1} e^{-u-v} du dv \\ &\text{Let } u = f(z, t) = zt \text{ and } v = g(z, t) = z(1-t) \\ &J(z, t) = \left| \begin{array}{cc} ab & z \\ (1-t) & -z \end{array} \right| = -z \end{aligned}$$

$$\begin{aligned} & \Gamma(a)\Gamma(b) \\ &= \int_{z=0}^{\infty} \int_{t=0}^1 (zt)^{a-1} [z(1-t)]^{b-1} e^{-z} |J(z, t)| dt dz \\ &= \int_{z=0}^{\infty} \int_{t=0}^1 (zt)^{a-1} [z(1-t)]^{b-1} e^{-z} z dt dz \\ &= \int_{z=0}^{\infty} z^{a+b-1} e^{-z} \int_{t=0}^1 t^{a-1} (1-t)^{b-1} dt \\ &= \Gamma(a+b) B(a, b) \\ &\therefore B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} \end{aligned}$$

0.11 Location and Scale Parameters

In each of the following definitions, $F_0(z)$ represents a completely specified CDF, and $f_0(z)$ is the pdf.

Definition 11. Location Parameters A quantity η is a location parameter for the distribution of X if the CDF has the form

$$F(x) = F_0(x - \eta)$$

In other words, the pdf has the form

$$f(x) = f_0(x - \eta)$$

For example,
Consider the pdf

$$f_0(z) = e^{-|z|}, -\infty < z < \infty$$

If X has pdf of the form

$$f(x) = e^{-|x-\eta|}, -\infty < x < \infty$$

then η is the location parameter.

Definition 13. Location-Scale Parameter

Quantities η and $\theta > 0$ are called location-scale parameters for the distribution of X if the CDF has the form

$$F(x) = F_0\left(\frac{x - \eta}{\theta}\right)$$

In other words, the pdf has the form

$$f(x) = f_0\left(\frac{x - \eta}{\theta}\right)$$

The normal distribution is the most commonly encountered location-scale distribution, but there are other important examples.

$$X \sim CAU(\theta, \eta)$$

$$f(x) = \frac{1}{\theta\pi\left[1 + \left(\frac{x - \eta}{\theta}\right)^2\right]} \quad -\infty < x < \infty$$

0.11.2 Two-parameter Exponential Distribution

Another location-scale distribution, which is frequently encountered in life testing applications, has pdf

$$f(x) = \frac{1}{\theta} e^{-\left(\frac{x-\eta}{\theta}\right)} \quad x > \eta$$

and zero otherwise. This is called the **two-parameter exponential distribution**, denoted by

$$X \sim EXP(\eta, \theta)$$

The mean, variance and MGF of an Exponential distribution are:

$$E(X) = \eta + \theta, \sigma^2 = \theta^2, M_X(t) = \left(\frac{e^{\eta t}}{1 - \theta t} \right)$$

0.11.3 Double-Exponential Distribution

If X has pdf of the form

$$f(x) = \frac{1}{2\theta} e^{-|x-\eta|/\theta} \quad -\infty < x < \infty$$

and zero otherwise.

This location-scale distribution is called the **Laplace** or **double-exponential** distribution, denoted by

$$X \sim DE(\theta, \eta)$$

The mean, variance and MGF of an Exponential distribution are:

$$E(X) = \eta, \sigma^2 = 2\theta^2, M_X(t) = \left(\frac{e^{\eta t}}{1 - \theta^2 t^2} \right)$$

0.12 Joint Discrete Distributions

In many applications there will be more than one random variable of interest, say X_1, X_2, \dots, X_k .

It is convenient mathematically to regard these variables as components of a k -dimensional vector, $\mathbf{X} = (X_1, X_2, \dots, X_k)$, which is capable of

assuming values $x = (x_1, x_2, \dots, x_k)$ in a k -dimensional Euclidean space. Note, for example, that an observed value x may be the result of measuring k characteristics once each, or the result of measuring one characteristic k times.

Definition 14. The joint probability density function (joint pdf) of the k -dimensional discrete random variable $X = (X_1, X_2, \dots, X_k)$ is defined to be

$$f(x_1, x_2, \dots, x_k) = P[X_1 = x_1, X_2 = x_2, \dots, X_k = x_k]$$

for all possible values $x = (x_1, x_2, \dots, x_k)$ of X .

Theorem 17. A function $f(x_1, x_2, \dots, x_k)$ is the joint pdf for some vector-valued random variable

$$X = (x_1, x_2, \dots, x_k)$$

if and only if the following properties are satisfied

1. $f(x_1, x_2, \dots, x_k) > 0$ for all possible values x_1, x_2, \dots, x_k
2. $\sum_{x_1} \sum_{x_2} \cdots \sum_{x_k} f(x_1, x_2, \dots, x_k) = 1$

Example 23.

Let the joint pmf of X_1 and X_2 be defined by

$$p(x_1, x_2) = \frac{x_1 + x_2}{32}, \quad x_1 = 1, 2, x_2 = 1, 2, 3, 4.$$

- (a) Display the joint probability distribution of X_1 and X_2 in a table.

Sol:

		x_1	
		1	2
x_2	1	$\frac{1}{32}$	$\frac{3}{32}$
	2	$\frac{3}{32}$	$\frac{5}{32}$
x_2	3	$\frac{3}{32}$	$\frac{6}{32}$
	4	$\frac{3}{32}$	$\frac{6}{32}$

- (c) Find $P(X_1 < X_2)$.

Sol:

$$\begin{aligned} P(X_1 < X_2) &= p(1, 2) + p(1, 3) + p(1, 4) + p(2, 3) + p(2, 4) \\ &= \frac{3}{32} + \frac{4}{32} + \frac{5}{32} + \frac{6}{32} + \frac{6}{32} \\ &= \boxed{\frac{23}{32}} \end{aligned}$$

- (d) Find $P(X_1 + X_2 = 4)$.

Sol:

$$\begin{aligned} P(X_1 + X_2 = 4) &= p(1, 3) + p(2, 2) \\ &= \frac{4}{32} + \frac{4}{32} \\ &= \boxed{\frac{1}{4}} \end{aligned}$$

- (b) Verify that the probability function satisfies Theorem 17.

Sol:

(a) $p(x_1, x_2) \geq 0$ for all x_1 and x_2 .

$$\begin{aligned} (b) \sum_{x_1, x_2} p(x_1, x_2) &= \frac{2}{32} + \frac{3}{32} + \frac{3}{32} + \frac{4}{32} + \frac{4}{32} + \frac{5}{32} + \frac{5}{32} + \frac{6}{32} \\ &= 1 \end{aligned}$$

Definition 16. Joint CDF The joint cumulative distribution function of the k random variables X_1, X_2, \dots, X_k is the function defined by

$$F(x_1, x_2, \dots, x_k) = P[X_1 \leq x_1, \dots, X_k \leq x_k]$$

Theorem 18. A function $F(x_1, x_2)$ is a bivariate CDF if and only if

- $\lim_{x_1 \rightarrow -\infty} F(x_1, x_2) = F(-\infty, x_2) = 0 \quad \forall x_2$

- $\lim_{x_2 \rightarrow -\infty} F(x_1, x_2) = F(x_1, -\infty) = 0 \quad \forall x_1$

- $\lim_{x_1 \rightarrow \infty, x_2 \rightarrow \infty} F(x_1, x_2) = F(\infty, \infty) = 1 \quad \forall x_1, x_2$

- $F(b, d) - F(b, c) - F(a, d) + F(a, c) \geq 0 \quad \forall a < b, c < d$

- $\lim_{h \rightarrow 0^+} F(x_1 + h, x_2) = \lim_{h \rightarrow 0^+} F(x_1, x_2 + h) = F(x_1, x_2)$

Example 24.

A local supermarket has three checkout counters. Two customers arrive at the counters at different times when the counters are serving no other customers. Each customer chooses a counter at random, independently of the other. Let X_1 denote the number of customers who choose counter 1 and X_2 , the number who select counter 2.

- (a) Find the joint probability function of X_1 and X_2 .

Sol:

Let (i, j) pair denotes the event that the 1st customer choose counter i and the 2nd choose counter j , where $i, j = 1, 2, 3$.

$$S = \{(1, 1), (1, 2), (1, 3), (2, 1), (2, 2), (2, 3), (3, 1), (3, 2), (3, 3)\}$$

		x_2	
		0	1
	0	$\frac{1}{9}$	$\frac{2}{9}$
	1	$\frac{2}{9}$	$\frac{1}{9}$
2	$\frac{1}{9}$	0	0

(b) Find $F(-1, 2)$, $F(1.5, 2)$, and $F(5, 7)$.

Sol:

$$F(-1, 2) = P(X_1 \leq -1, X_2 \leq 2) = 0$$

$$F(1.5, 2) = P(X_1 \leq 1.5, X_2 \leq 2) = \frac{8}{9}$$

$$F(5, 7) = P(X_1 \leq 5, X_2 \leq 7) = 1$$

Example 25. If X and Y are discrete random variables with joint pdf

$$f(x, y) = c \frac{2^{x+y}}{x!y!} \quad x = 0, 1, 2, \dots; y = 0, 1, 2, \dots$$

and zero otherwise.

(a) Find the constant c .

Sol:

$$\begin{aligned} c \sum_{x=0}^{\infty} \sum_{y=0}^{\infty} \frac{2^x 2^y}{x! y!} &= 1 \\ c \sum_{x=0}^{\infty} \frac{2^x}{x!} \sum_{y=0}^{\infty} \frac{2^y}{y!} &= 1 \\ ce^2 e^2 &= 1 \\ c &= e^{-4} \end{aligned}$$

(b) Find the marginal pdf's of X and Y .

Sol:

$$\begin{aligned} f(x) &= \sum_{y=0}^{\infty} (e^{-4}) \frac{2^x}{x! y!} = (e^{-4}) \frac{2^x}{x!} \sum_{y=0}^{\infty} \frac{2^y}{y!} \\ &= (e^{-4}) \frac{2^x}{x!} e^2 \\ &= \frac{2^x e^{-2}}{x!} \end{aligned}$$

$$\text{Similarly, } f(y) = \frac{2^y e^{-2}}{y!}$$

0.13 Joint Continuous Distributions

Definition 17. A k -dimensional vector valued random variable $X = (X_1, X_2, \dots, X_k)$ is said to be continuous if there is a function $f(x_1, x_2, \dots, x_k)$, called the joint probability density function (joint pdf), of X , such that the joint CDF can be written as

$$F(x_1, x_2, \dots, x_k) = \int_{-\infty}^{x_k} \cdots \int_{-\infty}^{x_1} f(t_1, t_2, \dots, t_k) dt_1 \cdots dt_k \\ \forall x = (x_1, x_2, \dots, x_k).$$

Example 26.

Let X_1 denote the concentration of a certain substance in one trial of an experiment, and X_2 the concentration of the substance in a second trial of the experiment. Assume that the joint pdf is given by $f(x_1, x_2) = 4x_1 x_2; 0 < x_1 < 1, 0 < x_2 < 1$, and zero otherwise.

(a) Find the joint CDF.

Sol:

$$F(x_1, x_2) = \int_{-\infty}^{x_2} \int_{-\infty}^{x_1} f(t_1, t_2) dt_1 dt_2 \\ = \int_0^{x_2} \int_0^{x_1} 4t_1 t_2 dt_1 dt_2 \\ = 4 \int_0^{x_2} t_2 \left[\frac{t_1^2}{2} \right]_0^{x_1} dt_2 \\ = 2 \int_0^{x_2} t_2 x_1^2 dt_2 \\ = 2x_1^2 \left[\frac{t_2^2}{2} \right]_0^{x_2} \\ = x_1^2 x_2, 0 < x_1 < 1, 0 < x_2 < 1$$

Theorem 19. Any function $f(x_1, x_2, \dots, x_k)$ is a joint pdf of a k -dimensional random variable if and only if

1. $f(x_1, x_2, \dots, x_k) \geq 0 \quad \forall x_1, x_2, \dots, x_k$
2. $\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, x_2, \dots, x_k) dx_1 \cdots dx_k = 1$

$$F(x_1, x_2) = \int_0^{x_2} \int_0^1 f(t_1, t_2) dt_1 dt_2 \\ = \int_0^{x_2} \int_0^1 4t_1 t_2 dt_1 dt_2 \\ = 4 \int_0^{x_2} t_2 \left[\frac{t_1^2}{2} \right]_0^1 dt_2 \\ = 2 \int_0^{x_2} t_2 dt_2 \\ = 2 \left[\frac{t_2^2}{2} \right]_0^{x_2} \\ = x_2^2, x_1 > 1, 0 < x_2 < 1$$

$$\begin{aligned} F(x_1, x_2) &= \int_0^{x_1} \int_0^1 f(t_1, t_2) dt_1 dt_2 \\ &= \int_0^{x_1} \int_0^1 4t_1 t_2 dt_1 dt_2 \\ &= 4 \int_0^{x_1} t_2 [\frac{t_1^2}{2}]_0^1 dt_2 \\ &= 2 \int_0^{x_1} t_2 dt_2 \\ &= 2[\frac{t_2^2}{2}]_0^{x_1} \\ &= x_1^2, x_2 > 1, 0 < x_1 < 1 \end{aligned}$$

$$F(x_1, x_2) = \int_0^\infty \int_0^\infty 4t_1 t_2 dt_1 dt_2 = 1, 0 < x_1, x_2 < \infty$$

(b) Find $P\left[\frac{X_1+X_2}{2} < 0.5\right]$.

Sol:

$$\begin{aligned} P\left[\frac{X_1+X_2}{2} < 0.5\right] \\ &= P(X_1 + X_2 < 1) \\ &= \int_0^1 \int_{0-x_2}^1 4x_1 x_2 dx_1 dx_2 \\ &= \int_0^1 4x_2 [\frac{x_1^2}{2}]_0^{1-x_2} dx_2 \\ &= 2 \int_0^1 x_2 (1 - x_2)^2 dx_2 \\ &= 2 \int_0^1 (x_2 - 2x_2^2 + x_2^3) dx_2 \\ &= 2[\frac{x_2^2}{2} - \frac{2x_2^3}{3} + \frac{x_2^4}{4}]_0^1 \\ &= \frac{1}{6} \end{aligned}$$

Definition 18.

If $X = (X_1, X_2, \dots, X_k)$ is a k -dimensional random variable with joint CDF $F(x_1, x_2, \dots, x_k)$, then the marginal CDF of X is

$$F_j(x_j) = \lim_{x_i \rightarrow \infty, \text{all } i \neq j} F(x_1, \dots, x_j, \dots, x_k)$$

Furthermore, the marginal pdf is

$$f_j(x_j) = \int \cdots \int_{\text{all } i \neq j} f(x_1, \dots, x_j, \dots, x_k) dx_1 \dots dx_k$$

Example 27. Let X_1, X_2 , and X_3 be continuous with a joint pdf of the form $f(x_1, x_2, x_3) = c; 0 < x_1 < x_2 < x_3 < 1$, and zero otherwise, where c is a constant.

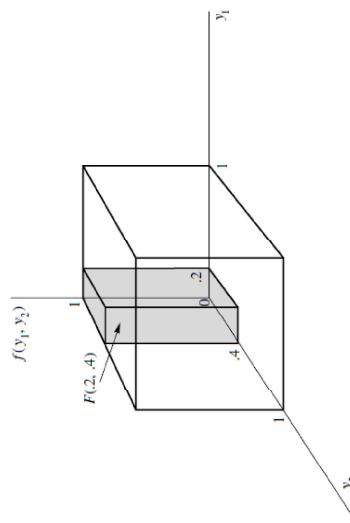
1. Find the marginal distributions of x_3 .
2. Find the joint pdf of the pair (X_1, X_2) .

Example 28.

Suppose that a radioactive particle is randomly located in a square with sides of unit length. That is, if two regions within the unit square and of equal area are considered, the particle is equally likely to be in either region. Let X_1 and X_2 denote the coordinates of the particle's location. A reasonable model for the relative frequency histogram for X_1 and X_2 is the bivariate analogue of the univariate uniform density function:

$$f(x_1, x_2) = \begin{cases} 1, & 0 \leq x_1 \leq 1, 0 \leq x_2 \leq 1, \\ 0, & \text{otherwise} \end{cases}$$

(a) Sketch the probability density surface.



(b) Find $F(2, .4)$.

Sol:

$$\begin{aligned} F(2, .4) &= P(X_1 \leq 2, X_2 \leq .4) \\ &= (.2)(.4) \\ &= \boxed{0.08} \end{aligned}$$

(c) Find $P(.1 \leq X_1 \leq .3, 0 \leq X_2 \leq .5)$

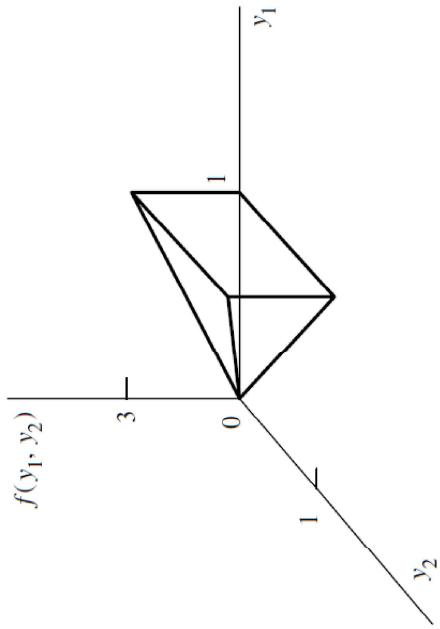
Sol:

$$\begin{aligned} P(.1 \leq X_1 \leq .3, 0 \leq X_2 \leq .5) &= 0.2(0.5) \\ &= \boxed{0.1} \end{aligned}$$

Example 29. The joint probability density function of X_1 and X_2 is

$$f(x_1, x_2) = \begin{cases} 3x_1, & 0 \leq x_2 \leq x_1 \leq 1, \\ 0, & \text{otherwise} \end{cases}$$

(a) Sketch the probability density surface.



(b) Find $P(0 \leq X_1 \leq .5, X_2 \geq 0.25)$.

$$\begin{aligned} Sol: \quad & P(0 \leq X_1 \leq .5, X_2 \geq 0.25) \\ & = \int_{.25}^{.5} \int_{.25}^{x_1} 3x_1 dx_2 dx_1 \\ & = \int_{.25}^{.5} 3x_1 [x_2]_{.25}^{x_1} dx_1 \\ & = \int_{.25}^{.5} 3x_1 (x_1 - .25) dx_1 \\ & = 3 \left[\frac{x_1^2}{2} - \frac{.25x_1}{2} \right]_{.25}^{.5} \\ & = \boxed{\frac{5}{128}} \end{aligned}$$

Theorem 20. If X_1 and X_2 are random variables with joint pdf $f(x_1, x_2)$ and marginal pdf's $f_1(x_1)$ and $f_2(x_2)$, then

$$f(x_1, x_2) = f_1(x_1)f(x_2|x_1) = f_2(x_2)f(x_1|x_2)$$

and if X_1 and X_2 are independent, then

$$f(x_2|x_1) = f_2(x_2)$$

and

$$f(x_1|x_2) = f_1(x_1)$$

Example 30.

Let

$$f(x_1, x_2, x_3, x_4) = \begin{cases} \frac{3}{4}(x_1^2 + x_2^2 + x_3^2 + x_4^2), & 0 < x_i < 1, i = 1, 2, 3, 4 \\ 0, & \text{otherwise} \end{cases}$$

(a) Find the marginal pdf of (X_1, X_2) .

- (b) Find the conditional pdf of (X_3, X_4) given $X_1 = \frac{1}{3}$ and $X_2 = \frac{2}{3}$.

Sol:

$$\begin{aligned} & f(x_1, x_2) \\ &= \int_0^\infty \int_0^\infty f(x, x_2, x_3, x_4) dx_3 dx_4 \\ &= \int_0^1 \int_0^1 \frac{3}{4}(x_1^2 + x_2^2 + x_3^2 + x_4^2) dx_3 dx_4 \\ &= \frac{3}{4}(x_1^2 + x_2^2) + \frac{1}{2} \\ & \text{for } 0 < x_1 < 1 \text{ and } 0 < x_2 < 1. \end{aligned}$$

$$\begin{aligned} & f(x_3, x_4|x_1, x_2) \\ &= \frac{f(x_1, x_2, x_3, x_4)}{f(x_1, x_2)} \\ &= \frac{\frac{3}{4}(x_1^2 + x_2^2 + x_3^2 + x_4^2)}{\frac{3}{4}(x_1^2 + x_2^2) + \frac{1}{2}} \\ &= \frac{x_3^2 + x_4^2}{x_1^2 + x_2^2 + \frac{2}{3}} \end{aligned}$$

$$\begin{aligned} & f(x_3, x_4|x_1 = \frac{1}{3}, x_2 = \frac{2}{3}) \\ &= \frac{(\frac{1}{3})^2 + (\frac{2}{3})^2 + x_3^2 + x_4^2}{(\frac{1}{3})^2 + (\frac{2}{3})^2 + \frac{2}{3}} \\ &= \frac{5}{11} + \frac{9}{11}x_3^2 + \frac{9}{11}x_4^2 \end{aligned}$$

Example 31. The joint density function of X_1 and X_2 is given by

$$f(x_1, x_2) = \begin{cases} 30x_1x_2^2, & x_1 - 1 \leq x_2 \leq 1 - x_1, 0 \leq x_1 \leq 1 \\ 0, & \text{otherwise} \end{cases}$$

(a) Show that the marginal density of X_1 is a beta density with $a = 2$ and $b = 4$.

Sol:

$$\begin{aligned} f_1(x_1) &= \int_{x_1-1}^{1-x_1} 30x_1x_2^2 dx_2 \\ &= 30x_1 \left[\frac{x_2^3}{3} \right]_{x_1-1}^{1-x_1} \\ &= 10x_1[(1-x_1)^3 + (1-x_1)^3] \\ &= 20x_1(1-x_1)^3, 0 \leq x_1 \leq 1 \end{aligned}$$

(b) Derive the marginal density of X_2 .

Sol:

$$\begin{aligned} \text{For } 0 < x_2 < 1 \\ f_2(x_2) &= \int_0^{1-x_2} 30x_1x_2^2 dx_1 \\ &= 30x_2^2 \left[\frac{x_1^2}{2} \right]_0^{1-x_2} \end{aligned}$$

$$= 15x_2^2(1-x_2)^2, 0 < x_2 < 1$$

$$\begin{aligned} \text{For } -1 < x_2 < 0 \\ f_2(x_2) &= \int_0^{1+x_2} 30x_1x_2^2 dx_1 \\ &= \int_0^{1+x_2} 30x_1x_2^2 dx_1 \\ &= 30x_2^2 \left[\frac{x_1^2}{2} \right]_0^{1+x_2} \\ &= 15x_2^2(1+x_2)^2, -1 < x_2 < 0 \end{aligned}$$

$$\begin{aligned} f(x_2) &= \begin{cases} 15x_2^2(1-x_2)^2, & 0 < x_2 < 1 \\ 15x_2^2(1+x_2)^2, & -1 < x_2 < 0 \end{cases} \end{aligned}$$

(c) Derive the conditional density of X_2 given $X_1 = x_1$.

Sol:

$$\begin{aligned} f(x_2|x_1) &= \frac{f(x_2)}{f(x_1|x_2)} \\ &= \frac{f_2(x_2)}{f_1(x_1)} \\ &= \frac{30x_1x_2^2}{20x_1(1-x_1)^3} \end{aligned}$$

$$= \frac{3x_2^2}{2(1-x_1)^3}, x_1 - 1 \leq x_2 \leq 1 - x_1$$

(d) Find $P(X_2 > 0 | X_1 = .75)$.

Sol:

$$\begin{aligned} f(x_2 | x_1 = 0.75) \\ &= \frac{3x_2^2}{2(1-.75)^3}, 0.75 - 1 \leq x_2 \leq 1 - 0.75 \\ &= 96x_2^2, -.25 \leq x_2 \leq .25 \end{aligned}$$

$$\begin{aligned} P(X_2 > 0 | X_1 = .75) \\ &= \int_{0.25}^{0.75} 96x_2^2 dx_2 \\ &= 96 \left[\frac{x_2^3}{3} \right]_0^{0.25} \\ &= 32[.25^3] \\ &= \boxed{0.5} \end{aligned}$$

0.15 Independent Random Variables

Definition 20. Independent Random Variables
Random variables X_1, \dots, X_k are said to
be independent if for every $a_i < b_i$,

$$\begin{aligned} P(a_1 \leq x_1 \leq b_1, \dots, a_k \leq x_k \leq b_k) \\ = \prod_{i=1}^k P(a_i \leq x_i \leq b_i) \end{aligned}$$

Theorem 21. Random variables X_1, \dots, X_k are
independent if and only if the following properties
holds:

$$\begin{aligned} F(x_1, \dots, x_k) &= F_1(x_1) \cdots F_k(x_k) \\ f(x_1, \dots, x_k) &= f_1(x_1) \cdots f_k(x_k) \end{aligned}$$

where $F_i(x_i)$ and $f_i(x_i)$ are the marginal CDF
and pdf of X_i , respectively,

Theorem 22. Two random variables X_1 and X_2 with joint pdf $f(x_1, x_2)$ are independent if and only if:

1. The “support set” $\{(x_1, x_2) | f(x_1, x_2) > 0\}$, is a Cartesian product, $A \times B$, and
2. The joint pdf can be factored into the product of functions of x_1 and x_2 , $f(x_1, x_2) = g(x_1)h(x_2)$

Example 32. The joint pdf of a pair X_1 and X_2 is

$$f(x_1, x_2) = 8x_1x_2, 0 < x_1 < x_2 < 1$$

and zero otherwise. Are X_1 and X_2 independent?

Sol: This function can clearly be factored according to part (2) of the theorem, but the support set, $\{(x_1, x_2), 0 < x_1 < x_2 < 1\}$, is a triangular region that cannot be represented as a Cartesian product. Thus, X_1 and X_2 are dependent.

Example 33. Consider now a pair X_1 and X_2 with joint pdf

$$f(x_1, x_2) = x_1 + x_2, 0 < x_1 < 1, 0 < x_2 < 1$$

and zero otherwise. Are X_1 and X_2 independent?

Sol: In this case the support set is $\{(x_1, x_2), 0 < x_1 < 1 \text{ and } 0 < x_2 < 1\}$, which can be represented as $A \times B$, where A and B are both the open interval $(0, 1)$. However, part (2) of the theorem is not satisfied because $x_1 + x_2$ cannot be factored as $g(x_1)h(x_2)$. Thus, X_1 and X_2 are dependent

Example 34. The joint distribution of X_1 and X_2 is given by the entries in the following table.

	x_2	
x_1	0	1
0	0.12	0.28
1	0.18	0.42

Show that X_1 and X_2 are independent.

Sol:

$$\begin{aligned}
 p_1(0) &= 0.12 + 0.28 = 0.4, \quad p_1(1) = .18 + .42 = .6 \\
 p_2(0) &= 0.12 + 0.18 = 0.3, \quad p_2(1) = .28 + .42 = .7 \\
 p_1(0)p_2(0) &= .4(.3) = .12 = p(0, 0) \\
 p_1(0)p_2(1) &= .4(.7) = .28 = p(0, 1) \\
 p_1(1)p_2(0) &= .6(.3) = .18p(1, 0) \\
 p_1(1)p_2(1) &= .6(.7) = .42 = p(1, 1)
 \end{aligned}$$

Is X_1 independent of X_2 ?

Sol:

$$\begin{aligned}
 p_1(0) &= 4/9; \quad p_2(0) = 4/9 \\
 p_1(0)p_2(0) &= 16/81 \neq p(0, 0). \quad \text{Thus } X_1 \text{ and } X_2 \\
 &\text{are not independent.}
 \end{aligned}$$

Example 35.

The joint distribution of X_1 and X_2 is given by the entries in the following table.

	x_2		
x_1	0	1	2
0	1/9	2/9	1/9
1	2/9	2/9	0
2	1/9	0	0

Is X_1 independent of X_2 ?

Example 36. Let

$$f(x_1, x_2) = \begin{cases} 2, & 0 \leq x_2 \leq x_1 \leq 1 \\ 0, & \text{otherwise} \end{cases}$$

Show that X_1 and X_2 are dependent.

Sol:

$$f_1(x_1) = \int_0^{x_1} 2dx_2 = 2x_1, 0 \leq x_1 \leq 1 \quad \text{otherwise}$$

$$f_2(x_2) = \int_{x_2}^1 2dx_1 = 2(1 - x_2), 0 \leq x_2 \leq 1 \quad \text{otherwise}$$

$$f_1(x_1)f_2(x_2) = 4x_1(1 - x_2) \neq f(x_1x_2)$$

Thus X_1 and X_2 are dependent.

0.16 The Expected Value of a Function of Random Variables

Definition 21. If $X = (X_1, \dots, X_k)$ has a joint pdf $f(x_1, \dots, x_k)$, and if $Y = u(X_1, \dots, X_k)$ is a function of X , then $E(Y) = E[u(X_1, \dots, X_k)]$, where

$$E_X[u(X_1, \dots, X_k)] = \sum_{x_1} \cdots \sum_{x_k} u(x_1, \dots, x_k) f(x_1, \dots, x_k)$$

if X is discrete, and

$$\begin{aligned} E_X[u(X_1, \dots, X_k)] &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} u(x_1, \dots, x_k) f(x_1, \dots, x_k) dx_1 \cdots dx_k \\ &\quad \text{if } X \text{ is continuous.} \end{aligned}$$

Theorem 23. If X_1 and X_2 are random variables with joint pdf $f(x_1, x_2)$, then

$$E(X_1 + X_2) = E(X_1) + E(X_2)$$

Sol: We will show this for the continuous case:

$$\begin{aligned} E(X_1 + X_2) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_1 + x_2) f(x_1, x_2) dx_1 dx_2 \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_1) f(x_1, x_2) dx_1 dx_2 \\ &\quad + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_2) f(x_1, x_2) dx_1 dx_2 \\ &= \int_{-\infty}^{\infty} (x_1) \int_{-\infty}^{\infty} f(x_1, x_2) dx_2 dx_1 \\ &\quad + \int_{-\infty}^{\infty} (x_2) \int_{-\infty}^{\infty} f(x_1, x_2) dx_1 dx_2 \\ &= \int_{-\infty}^{\infty} (x_1) f(x_1) dx_1 + \int_{-\infty}^{\infty} (x_2) f(x_2) dx_2 \\ &= E(X_1) + E(X_2) \end{aligned}$$

Theorem 24. If X and Y are independent random variables and $g(x)$ and $h(y)$ are functions, then

$$E[g(X)h(Y)] = E[g(X)]E[h(Y)]$$

Sol:

$$\begin{aligned} E[g(X)h(Y)] &= E[g(X)]E[h(Y)] \\ &= \int \int g(x)h(y)f(x, y)dx dy \\ &= \int \int g(x)h(y)f(x)f(y)dx dy \\ &= \int g(x)f(x)dx \int h(y)f(y)dy = E[g(X)]E[h(Y)] \end{aligned}$$

It is possible to generalize this theorem to more than two variables. Specifically, if X_1, X_2, \dots, X_k are independent random variables, and

$u_1(x_1), \dots, u_k(x_k)$ are functions, then

$$E[u_1(X_1) \cdots u_k(X_k)] = E[u_1(X_1)] \cdots E[u_k(X_k)]$$

It is possible to combine the preceding theorems to show that if a_1, a_2, \dots, a_k , are constants and X_1, X_2, \dots, X_k are jointly distributed random variables, then

$$E\left(\sum_{i=1}^k a_i X_i\right) = \sum_{i=1}^k a_i E(X_i)$$

Example 37.

The joint distribution of X_1 and X_2 is given by the entries in the following table.

	x_2	
x_1	0	1
0	$\frac{1}{9}$	$\frac{2}{9}$
1	$\frac{2}{9}$	0
2	$\frac{1}{9}$	0

(a) Find $E(X_1)$

Sol:

$$E(X_1) = \sum_{x_1} \sum_{x_2} x_1 p(x_1, x_2) = 0(1/9) + 1(2/9) + 2(1/9) + 0(2/9) + 1(2/9) + 2(0/9) + 1(0) + 2(0) = \frac{2}{3}$$

(b) Find $V(X_1)$

Sol:

$$E(X_1^2) = \sum_{x_1} \sum_{x_2} x_1^2 p(x_1, x_2) = 0(1/9) + 1(2/9) + 4(1/9) + 0(2/9) + 1(2/9) + 1(0) + 4(0) = \frac{8}{9}$$

$$\begin{aligned} 0(1/9) + 1(0) + 4(0) &= \frac{8}{9} \\ V(X_1) &= E(X_1^2) - E^2(X_1) = \frac{8}{9} - (\frac{2}{3})^2 = \frac{4}{9} \end{aligned}$$

(c) Find $E(X_1 X_2)$

Sol:

$$\begin{aligned} E(X_1 X_2) &= \sum_{x_1} \sum_{x_2} x_1 x_2 p(x_1, x_2) = 0(0)(1/9) + \\ &1(0)(2/9) + 2(0)(1/9) + 0(1)(2/9) + (1)(1)(2/9) + \\ &(1)(2)(0/9) + (2)(0)(1/9) + 2(1)(0) + 2(2)(0) \\ &= \frac{2}{9} \end{aligned}$$

Example 38.

Let $f(x_1, x_2) = \begin{cases} 2x_1, & 0 \leq x_1 \leq 1; 0 \leq x_2 \leq 1 \\ 0, & \text{otherwise} \end{cases}$

(a) Find $E(X_1)$

Sol:

$$\begin{aligned} E(X_1) &= \int_0^1 \int_0^1 2x_1^2 dx_2 dx_1 = \int_0^1 2x_1^2 x_2]_0^1 dx_1 \\ &= \int_0^1 2x_1^2 dx_1 = \frac{2x_1^3}{3}]_0^1 = \frac{2}{3} \end{aligned}$$

(b) Find $V(X_1)$

Sol:

$$\begin{aligned} E(X_1^2) &= \int_0^1 \int_0^1 2x_1^3 dx_2 dx_1 = \int_0^1 2x_1^3 x_2]_0^1 dx_1 \\ &= \int_0^1 2x_1^3 dx_1 = \frac{2x_1^4}{4}]_0^1 = \frac{1}{2} \end{aligned}$$

$$V(X_1) = \frac{1}{2} - \frac{4}{9} = \frac{1}{18}$$

(c) Find $E(X_1^2 X_2)$

Sol:

$$\begin{aligned} E(X_1^2 X_2) &= \int_0^1 \int_0^1 2x_1^3 x_2 dx_2 dx_1 = \int_0^1 x_1^3 x_2]_0^1 dx_1 \\ &= \int_0^1 x_1^3 dx_1 = \frac{x_1^4}{4}]_0^1 = \frac{1}{4} \end{aligned}$$

Example 40.

Suppose X_1 and X_2 are independent random variables, $E(X_1) = 2$ and $E(X_2) = \frac{1}{3}$. Find $E(X_1 X_2)$.

Sol:

$$\begin{aligned} E(X_1 X_2) &= E(X_1)E(X_2) \text{ Since } X_1 \text{ and } X_2 \text{ are independent random variables} \\ &= 2\left(\frac{1}{3}\right) \\ &= \frac{2}{3} \end{aligned}$$

Definition 22. The covariance of a pair of random variables X and Y is defined by

$$Cov(X, Y) = E[(X - \mu_X)(Y - \mu_Y)]$$

Another common notation for covariance is σ_{XY} .

Theorem 25. If X and Y are random variables and a and b are constants, then

- $Cov(aX, bY) = abCov(X, Y)$
- $Cov(X + a, Y + b) = Cov(X, Y)$
- $Cov(X, aX + b) = aV(X)$

Sol:

- $Cov(aX, bY)$
 $= E(aX - aE(X))(bY - bE(Y))$
 $= aE[X - E(X)]b[Y - E(Y)]$
 $= abE[X - E(X)][Y - E(Y)]$
 $= abCov(X, Y)$
- $Cov(X + a, Y + b)$
 $= E[X + a - E(X + a)][(Y + b - E(Y + b))]$

$$\begin{aligned} &= E[X + a - E(X) - a][Y + b - E(Y) - b] \\ &= E[X - E(X)][Y - E(Y)] \\ &= Cov(X, Y) \end{aligned}$$

$$\begin{aligned} &\bullet Cov(X, aX + b) \\ &= E[X - E(X)][aX + b - E(aX + b)] \\ &= E[X - E(X)][aX + b - E(aX) - b] \\ &= E[X - E(X)][aX - E(aX)] \\ &= E[X - E(X)][a(X - E(X))] \\ &= aE[X - E(X)][X - E(X)] \\ &= aV(X) \end{aligned}$$

Theorem 26. If X and Y are random variables, then

$$Cov(X, Y) = E(XY) - E(X)E(Y)$$

and $Cov(X, Y) = 0$ whenever X and Y are independent.

Sol:

$$\begin{aligned} Cov(X, Y) &= E[X - E(X)][Y - E(Y)] \\ &= E[XY - XE(Y) - Y(E(X) + E(Y))] \\ &= E(XY) - E(X)E(Y) - EY)E(X) + . \\ &= E(XY) - E(X)E(Y) \end{aligned}$$

When X and Y are independent, $E(XY) = E(X)E(Y)$, hence $Cov(X, Y) = 0$.

Definition 23. If X and Y are random variables with variances σ_X^2 and σ_Y^2 and covariance $\sigma_{XY} = Cov(X, Y)$, then the correlation coefficient of X and Y is

$$\rho = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

The random variables X and Y are said to be uncorrelated if $\rho = 0$; otherwise they are said to be correlated.

Theorem 27. If ρ is the correlation coefficient of X and Y , then

$$-1 \leq \rho \leq 1$$

and $\rho = \pm 1$ if and only if $Y = aX + b$ with probability 1 for some $a \neq 0$ and b .

Theorem 28. If X_1 and X_2 are random variables with joint pdf $f(x_1, x_2)$, then

$$V(X_1 + X_2) = V(X_1) + V(X_2) + 2Cov(X_1, X_2)$$

and

$$V(X_1 + X_2) = V(X_1) + V(X_2)$$

whenever X_1 and X_2 are independent.

Sol:

$$\begin{aligned} V(X_1 + X_2) &= E[X_1 + X_2 - E(X_1 + X_2)]^2 \\ &= E[(X_1 + X_2)^2 - 2(X_1 + X_2)E(X_1 + X_2)] \\ &= E[(X_1 + X_2)^2 - 2E^2(X_1 + X_2) + E^2] \\ &= E[(X_1 + X_2)^2 - E^2(X_1 + X_2)] \\ &= E[X_1^2 + X_2^2 + 2X_1 X_2 - (E^2(X_1) + E^2(X_2))] \\ &= E[X_1^2 - E^2(X_1)] + E[X_2^2 - E^2(X_2)] \\ &= V(X_1) + V(X_2) + 2Cov(X_1, X_2) \end{aligned}$$

Whenever X_1 and X_2 are independent, $Cov(X_1, X_2) = 0$, hence, $V(X_1 + X_2) = V(X_1) + V(X_2)$.

It also can be verified that if a_1, a_2, \dots, a_k , are constants and X_1, X_2, \dots, X_k , are random variables, then

$$V\left(\sum_{i=1}^k a_i X_i\right) = \sum_{i=1}^k a_i^2 V(X_i) + 2 \sum_{i < k} \sum_{j < i} a_i a_j Cov(X_i, X_j)$$

and if X_1, X_2, \dots, X_k are independent, then

$$V\left(\sum_{i=1}^k a_i X_i\right) = \sum_{i=1}^k a_i^2 V(X_i)$$

Example 41. X_1 and X_2 have joint density given by

$$f(x_1, x_2) = \begin{cases} 2x_1, & 0 \leq x_1 \leq 1, 0 \leq x_2 \leq 1 \\ 0, & \text{otherwise} \end{cases}$$

Find $Cov(X_1, X_2)$.

Sol:

Since $f(x_1, x_2) = g(x_1)h(x_2)$ where $g(x_1) = y_1$ and $h(x_2) = y_2$. Then X_1 and X_2 are independent. Thus $Cov(X_1, X_2) = 0$

Example 42.

Let $f(x, y) = 6x$, $0 < x < y < 1$, and zero otherwise.
Find $Cov(X, Y)$.

Sol:

$$\begin{aligned}
 E(X) &= \int_0^1 \int_x^1 6x^2 dy dx = \int_0^1 6x^2(1-x) dx = 6 \left[\frac{x^3}{3} - \frac{x^4}{4} \right]_0^1 \\
 &= \frac{1}{2} \\
 E(Y) &= \int_0^1 \int_x^1 6xy dy dx = \int_0^1 6x[\frac{y^2}{2}]_x^1 dx = \int_0^1 6x[\frac{1}{2} - \frac{x^2}{2}] dx = \\
 &3 \left[\frac{x^2}{2} - \frac{x^4}{4} \right]_0^1 = \frac{3}{4} \\
 E(XY) &= \int_0^1 \int_x^1 6x^2 y dy dx = \int_0^1 6x^2 [\frac{y^2}{2}]_x^1 dx = \int_0^1 6x^2 [\frac{1}{2} - \\
 &\frac{x^2}{2}] dx = 3 \left[\frac{x^3}{3} - \frac{x^5}{5} \right]_0^1 = \frac{2}{5} \\
 Cov(X, Y) &= E(XY) - E(X)E(Y) = \frac{2}{5} - (\frac{1}{2})(\frac{3}{4}) = \boxed{\frac{1}{40}}
 \end{aligned}$$

Example 43.

Let X and Y be discrete random variables with joint pdf $f(x, y) = \frac{4}{5xy}$ if $x = 1, 2$ and $y = 2, 3$, and zero otherwise.
Find $Cov(X, Y)$.

Sol:

	y	
x	2	3
	$f_2(y)$	
1	$\frac{4}{10}$	$\frac{4}{15}$
2	$\frac{4}{20}$	$\frac{4}{30}$
	$f_1(x)$	$\frac{3}{5}$

$$\begin{aligned}
 E(X) &= 1(\frac{2}{3}) + 2(\frac{1}{3}) = \frac{4}{3} \\
 E(Y) &= 2(\frac{3}{5}) + 3(\frac{2}{5}) = 2.4 \\
 E(XY) &= 1(2)(\frac{4}{10}) + 1(3)(\frac{4}{15}) + 2(2)(\frac{4}{20}) + 2(3)(\frac{4}{30}) = 3.2 \\
 Cov(X, Y) &= E(XY) - E(X)E(Y) = 3.2 - (\frac{4}{3})(2.4) = 0
 \end{aligned}$$

Example 44. Let X_1 and X_2 be discrete random variables with joint probability distribution as shown in table below. Show that X_1 and X_2 are dependent but have zero covariance.

		x_2	
		-1	0
x_1	-1	0	1
-1	$\frac{1}{16}$	$\frac{3}{16}$	$\frac{1}{16}$
0	$\frac{3}{16}$	0	$\frac{3}{16}$
1	$\frac{1}{16}$	$\frac{3}{16}$	$\frac{1}{16}$

Sol:

		x_2		$p(x_1)$
		-1	0	
x_1	-1	$\frac{1}{16}$	$\frac{3}{16}$	$\frac{1}{16}$
-1	$\frac{3}{16}$	0	$\frac{3}{16}$	$\frac{6}{16}$
0	$\frac{1}{16}$	$\frac{3}{16}$	$\frac{1}{16}$	$\frac{3}{16}$
1	$\frac{1}{16}$	$\frac{3}{16}$	$\frac{1}{16}$	$\frac{3}{16}$
$p(x_2)$	$\frac{5}{16}$	$\frac{6}{16}$	$\frac{5}{16}$	

$$p_1(-1)p_2(-1) = \left(\frac{5}{16}\right)^2 = \frac{25}{256} \neq p(-1, -1)$$

Thus X_1 and X_2 are dependent.

$$E(X_1) = (-1)\left(\frac{5}{16}\right) + 0\left(\frac{6}{16}\right) + (1)\left(\frac{1}{16}\right) = 0 = E(X_2)$$

$$\begin{aligned} E(X_1 X_2) &= (-1)\left(-1\right)\left(\frac{1}{16}\right) + (-1)(0)\left(\frac{3}{16}\right) + (-1)(1)\left(\frac{1}{16}\right) + \\ &(0)(-1)\left(\frac{3}{16}\right) + (0)(0)\left(0\right) + (0)(1)\left(\frac{3}{16}\right) + (1)(-1)\left(\frac{1}{16}\right) + \\ &(1)(1)\left(\frac{1}{16}\right) = 0 \end{aligned}$$

$$Cov(X_1, X_2) = 0$$